Berkala Ilmiah Pendidikan
Volume 2 Nomor 2, Juli 2022

scidac plus

EVEKTIFITAS INSTRUMEN PENILAIAN PENDIDIKAN AGAMA ISLAM UNTUK MENINGKATKAN KEMAMPUAN BERPIKIR KRITIS DAN KREATIF PESERTA DIDIK DI SMK MUHAMMADIYAH MLATI YOGYAKARTA

Moh. Solikul Hadi[1*], Ahmad Zarnuji[2], M. Ikhwanuddin[3], Hasyim Asy'ari[4]
[1] Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia
[234] Institut Agama Islam Ma'arif NU (IAIMNU) Metro Lampung, Indonesia
mohsolikulhadi@gmail.com*

Abstrak

Mengingat kompetensi anak-anak Indonesia memang masih rendah dibandingkan negara-negara lainnya. Berdasarkan data Kemendikbud melalui Asesmen Kompetensi Peserta didik Indonesia (AKSI) 2016 ditemukan bahwa secara nasional 73,61% pencapaian kompetensi peserta didik masih berada pada posisi kurang. Kemendikbud mulai memberlakukan 10% soal yang membutuhkan daya nalar tingkat tinggi (high order thinking skills/HOTS). Meskipun baru 10%, ternyata banyak keluhan dari anak- anak Indonesia. Padahal melihat pembelajaran pada era abad ke 21 menggunakan dan mengandung muatan Communication Skill, Collaboration skill, critical thinking and Problem solving skill, creativity and innovation skill atau sering disebut dengan muatan 4C. Hasil penelitian ini menunjukkan bahwa kemampuan guru dalam menyusun instrumen penilaian belum sesuai dengan kompetensi yang diinginkan pada jenjang SMA/MA. Pada beberapa soal masih terdapat instrumen yang memiliki kategori rendah/ low order thinking, hasil analisis perhitungan data, tidak terdapat butir soal yang berkualitas sangat baik, 2 butir soal berkualitas baik, 16 butir soal berkualitas sedang, 16 soal berkualitas tidak baik, dan terdapat 6 butir soal yang memiliki kualitas sangat tidak baik, Kurang efektif karena melihat banyaknya instrumen butir soal tidak baik dan mesti dibuang serta melihat kata kata operasional yang digunakan dalam penyusunan instrument penilaian masih mengggunakan C1-C3. Maka soal atau instrumen yang digunakan belum efektif dalam meningkatkan daya kritis dan kreatif peserta didik.

Kata Kunci, Instrumen, Penilaian PAI, Berpikir kritis, Kreatif.

Abstract

Given the competence of Indonesian children is still low compared to other countries. Based on data from the Ministry of Education and Culture through the 2016 Indonesian Student Competency Assessment (AKSI) it was found that nationally 73.61% of the achievement of student competencies was still in a low position. The Ministry of Education and Culture has begun to apply 10% of questions that require high order thinking skills (HOTS). Even though it's only 10%, it turns out that there are many complaints from Indonesian children. In fact, seeing learning in the 21st century era uses and contains communication skills, collaboration skills, critical thinking and problem solving skills, creativity and innovation skills or often referred to as 4C

content. The results of this study indicate that the teacher's ability to develop assessment instruments is not in accordance with the desired competencies at the SMA/MA level. In some questions there are still instruments that have low order thinking categories, the results of data calculation analysis, there are no very good quality questions, 2 good quality questions, 16 medium quality questions, 16 poor quality questions, and there are 6 questions. Items that have very poor quality, Less effective because they see the number of instrument items that are not good and must be discarded and see the operational words used in the preparation of the assessment instrument still use C1-C3. So the questions or instruments used have not been effective in increasing the critical and creative power of students.

Keywords: *Instruments, PAI Assessment, Critical Thinking, Creative.*

PENDAHULUAN

Sikap daya kritis dan kreatif merupakan persoalan urgent dalam dunia pendidikan. Dunia pendidikan merupakan roda penggerak dalam mengawasi jalannya pemerintahan. Namun, dilihat akhir akhir ini, dunia pendidikan terbungkam dan terdiam tanpa suara. Tanpa terkecuali, dunia peserta didik lebih banyak diam dalam diskusi-diskusi kelas yang notabennya hampir rata-rata aktif dalam organisasi sekolah. Fenomena ini menjadi hal yang patut untuk bisa diuraikan dan diruntut akar persoalannya. Banyak faktor yang harus dianalisis secara tepat terkait timbulnya persoalan tersebut. Tak heran jika dilihat hasil kompetensi anak-anak Indonesia setiap tahunnya mengalami penurunan dibandingkan dengan negara yang lain. Menurut kabar yang disampaikan oleh Kementerian Pendidikan dan Kebudayaan (Kemendikbud) telah merilis pencapaian nilai *Programme for International Student Assessment* atau biasa disebut (PISA). Hasil PISA pada tahun 2018 yang dirilis bersama dengan 77 negara peserta survei PISA menunjukkan bahwa Peringkat siswa Indonesia mendudukkan di posisi 72 dari 77 negara *Organization for Economic Co-operation and Development* (OECD). Skor ini mengalami penurunan dari periode penilaian sebelumnya, yakni peringkat 62 dari 70 negara di 2015. Krisis pembelajaran ini juga ditemukan di beberapa negara yang berpenghasilan rendah, yaitu Afrika Sub Sahara sebesar 87 persen, Asia Timur Tengah sebesar 63 persen, Asia Selatan sebesar 58 persen, Amerika Latin dan Caribbean sebesar 51 persen, Asia Timur dan Pasific sebesar 21 persen, serta Eropa dan Asia Tengah sebesar 13 persen (Https://Republika.Co.Id/).

Disamping itu, data Asesmen Kompetensi Peserta didik Indonesia (AKSI) Kemendikbud ditemukan bahwa secara nasional pencapaian kompetensi peserta didik masih berada pada taraf yang kurang. Padahal jika dicermati, Kemendikbud sudah memberlakukan 10% soal yang membutuhkan daya nalar tingkat tinggi (*high order thinking skills*/HOTS). Meskipun hanya 10% ternyata massih banyak keluhan dari peserta didik di Indonesia. Jika dilihat pembelajaran era abad ke 21 peserta didik harus mampu bersaing dan menjadi pribadi yang unggul dengan dibekali berbagai skill dari guru, maka tidak heran jika era abad ke 21 guru dalam proses pembelajaran menggunakan dan mengandung muatan *Communication Skill, Collaboration skill, critical thinking and Problem solving skill, creativity and innovation skill* atau sering disebut dengan muatan 4C (Mispani et al., 2021).

Kemampuan berpikir, baik berpikir kritis maupun berpikir kreatif merupakan kemampuan dasar yang harus dimiliki oleh peserta didik agar mampu memecahkan persoalan kehidupan yang senantiasa berubah. Dengan demikian, pengembangan kemampuan kritis dan berpikir kreatif menjadi hal yang perlu dilakukan dan dilatih sejak jenjang dasar sampai jenjang pendidikan menengah. Trianto mengatakan bahwa kemampuan berpikir kritis dan kreatif peserta didik dapat dilatih dengan pembelajaran yang menuntut peserta didik untuk melakukan eksplorasi, inkuiri, penemuan dan memecahkan masalah serta melalui belajar dalam kelompok kecil dengan menerapkan pendekatan *scaffolding* yang nantinya mampu memberikan stimulun

peningkatan daya kritis dan kreatif, Sehingga pada dasarnya selama pembelajaran, peserta didik dituntut untuk lebih aktif (Marniwati, 2019).

Kemampuan peserta didik dalam menjawab soal HOTS sangat diperlukan, dengan melihat persaingan dunia kerja semakin lebih kompleks. Kemampuan menjawab soal HOTS yang nantinya mampu meningkatkan daya kritis dan kreatif peserta didik menjadi salah satu dasar pemikiran peserta didik untuk mengantarkan masa deoan yang cerah. Peserta didik dihadapkan mampu menjawab persoalan-persoalan yang terjadi di masyarakat dengan bijak dan tepat. Jika tidak ada semangat untuk mengasah daya kritis dan kreatif pada diri peserta didik, maka selamanya pendidikan Indonesia akan menjadi peringkat bawah dimata dunia. Pengalaman proses pendidikan selama setengah abad menunjukkan bahwa setiap kali muncul pembahasan yang mengarah kepada upaya perbaikan sistem pendidikan nasional, selalu yang menjadi titik berat perhatian adalah pembenahan kurikulum dan sumber daya mansuianya yaitu pendidik dan peserta didik.

METODE

Penelitian ini merupakan penelitian kuantitatif (Sugiono, 2014) dengan subyek penelitian adalah peserta didik SMK Muhammdiyah Mlati. Sumber data didapat dari kepala sekolah, Waka Kurikulum, Koordinator ISMUBA serta Guru ISMUBA yang ada di SMK Muhammadiyah Mlati. Teknik pengumpulan data dilakukan dengan teknik pengumpulan data observasi, wawancara dan olah data. Setelah semua data didapatkan kemudian dianalisis dengan teknik analisis kualitas instrumen uji validitas dengan menggunakan Korelasi Product Moment, Uji Reliabilitas, Tingkat kesukaran, daya pembeda, serta efektivitas pengecoh soal/instrument (Suliyanto, 2017).

HASIL DAN PEMBAHASAN

Melihat soal apakah valid dan tidaknya, maka harus ada penilaian atau pengujian. Pengujian berfungsi sebagai penentu valid atau tidaknya butir pernyataan dengan melakukan perbandingan harga *r* hitung *product moment* dan kemudian dikonsultasikan dengan harga *r* tabel *product moment* dengan kriteria apabila harga *r* hitung sama atau lebih besar dengan harga *r* tabel maka soal dinyatakan valid. Melihat harga *r* tabel perlu dicari terlebih dahulu derajat kebebasan (*degree of freedom*) atau disingkat df dengan rumus: df= N-nr, dimana N adalah banyaknya peserta tes (*testee)* dan nr adalah banyaknya variabel yang dikorelasikan (Alwi, 2015). Maka dalam hal ini, banyak butir soal (N) adalah 40, sedang banyak variabel yang dikorelasikan (nr) adalah 2. Dengan demikian df= 40-2 = 38, sedangkan taraf signifikansi untuk uji validitas yang diapakai umumnya 5%. Dengan df=38 dan taraf signifikansi 5% diperoleh *r* tabel sebesar 0,320 (Sukiman, 2012). Selanjutnya dibandingkan antara harga *r* hitung dengan *r* tabel. Hasil uji validitas instrumen dapat dilihat pada tabel berikut ini:

Variabel	r hitung	r table	Keterangan
Soal 1	0.117	0.330	Tidak valid
Soal 2	0.225	0.340	Tidak valid
Soal 3	0.346	0.310	Valid
Soal 4	0.232	0.340	Tidak Valid
Soal 5	0.246	0.340	Tidak Valid
Soal 6	0.467	0.360	Valid
Soal 7	0.245	0.360	Tidak Valid
Soal 8	0.117	0.320	Tidak Valid
Soal 9	0.25	0.320	Tidak Valid
Soal 10	0.336	0.320	Valid
Soal 11	0.233	0.320	Tidak valid

Soal 12	0.46	0.320	Valid
Soal 13	0.108	0.320	Tidak valid
Soal 14	0.346	0.320	Valid
Soal 15	0.311	0.320	Tidak valid
Soal 16	0.569	0.320	Valid
Soal 17	0.147	0.320	Tidak valid
Soal 18	0.330	0.320	Valid
Soal 19	0.426	0.320	Valid
Soal 20	0.224	0.310	Tidak valid
Soal 21	0.147	0.310	Tidak valid
Soal 22	0.245	0.310	Tidak valid
Soal 23	0.344	0.320	Valid
Soal 24	0.427	0.320	Valid
Soal 25	0.106	0.310	Tidak valid
Soal 26	0.376	0.330	Valid
Soal 27	0.232	0.360	Tidak valid
Soal 28	0.246	0.360	Tidak valid
Soal 29	0.467	0.320	Valid
Soal 30	0.108	0.350	Tidak valid
Soal 31	0.346	0.320	Valid
Soal 32	0.311	0.320	Tidak valid
Soal 33	0.569	0.320	Valid
Soal 34	0.232	0.310	Tidak valid
Soal 35	0.246	0.310	Tidak valid
Soal 36	0.330	0.320	Valid
Soal 37	0.226	0.320	Tidak valid
Soal 38	0.246	0.320	Tidak valid
Soal 39	0.232	0.210	Tidak valid
Soal 40	0.265	0.230	Tidak valid

Hasil validasi soal PAS semester 1 Pembelajaran ISMUBA Tahun ajaran 2019/2020 terdapat 40 soal (Solichin, 2017). Berdasarkan hasil penghitungan diatas 15 soal dinyatakan valid dengan prosentase sebesar 40% yakni nomor soal 3, 6 ,10, 12, 14, 16, 18, 19, 23, 24, 26, 29, 31, 33, 36, Kemudian terdapat 25 soal dinyatakan tidak valid dengan prosentase sebesar 60% yakni 1, 2, 4, 5, 7, 8, 9, 11, 13, 15, 17, 20, 21, 22, 25, 27, 28, 30, 32,34, 35, 37, 38, 39, 40.

Pengujian reliabilitas Hasil validasi soal PAS semester 1 Pembelajaran ISMUBA Tahun ajaran 2019/2020 dengan menggunakan SPSS (Hartini & Martin, 2020). Setelah didapatkan hasil perhitungan, diketahui soal tersebut memiliki reliabilitas sebesar 0, 634, sehingga dapat disimpulkan bahwa soal PAS semester gasal Pembelajaran ISMUBA Tahun ajaran 2019/2020 memiliki tingkat reliabilitas yang rendah.

Perhitungan indeks kesukaran dilakukan dengan memberikan interpretasi terhadap hasil perhitungan (Fatimah & Alfath, 2019). Cara memberikan interpretasi adalah dengan dikonsultasikan hasil perhitungan indeks tingkat kesukaran tersebut dengan suatu patokan atau kriteria sebagai berikut soal dengan P 0,00 sampai 0,30 adalah soal yang tergolong sukar, soal dengan P 0,31 sampai 0,70 adalah soal yang tergolong sedang; dan soal dengan P 0,71 sampai 1,00 adalah soal yang tergolong mudah. Berdasarkan hasil perhitungan terdapat 14 butir soal atau sebesar 32% yang termasuk dalam soal yang sukar, 16 soal atau sebesar 43% tergolong dalam soal sedang, dan sebanyak 10 soal atau 25% termasuk dalam soal yang mudah. Berikut penjabaran butir soal bersdasarkan

tingkat atau indeks kesukaran.

Pada perhitungan daya pembeda menggunakan bantuan program *Excel*. Jumlah subyek dalam penelitian ini adalah 80 siswa, sehingga termasuk kelompok kecil. Pada kelompok kecil untuk menghitung daya pembeda terlebih dahulu dibagi dalam 50% kelompok atas dan 50% kelompok bawah. Hasil perhitungan daya pembeda diinterpretasikan ke dalam lima kriteria yaitu:

D : 0,00 – 0,20 maka daya pembeda butir soal tersebut jelek

D : 0,21 – 0,40 berarti daya pembeda soal cukup

D : 0,41 – 0,70 berarti daya pembeda soal tersebut baik

D : 0,71 – 1,00 berarti daya pembeda butir soal tersebut baik sekali

D : negatif berarti daya pembeda soal tersebut tidak baik dan dibuang.

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Daya pembeda dihitung dengan rumus sebagai berikut:

Jumlah kelompok atas terdiri 24 siswa dan kelompok bawah ada 24 siswa. Cara mengetahui daya pembeda dengan mengurangkan hasil proporsi siswa yang menjawab benar pada kelompok atas dibagi banyaknya siswa kelompok atas dengan hasil proporsi siswa yang menjawab benar pada kelompok bawah dibagi banyaknya siswa kelompok bawah (Muluki, 2020). Berdasarkan hasil perhitungan daya pembeda dengan bantuan program *Excel* menunjukkan 14 butir soal atau sebesar 35% memiliki daya pembeda jelek, 14 soal atau sebesar 35% memiliki daya pembeda cukup, 4 butir soal atau sebesar 10% memiliki daya pembeda soal yang baik, sedangkan sebesar 8 soal yang bernilai negative sebesar 20% memiliki daya pembeda soal yang tidak baik dan harus dibuang. Berdasarkan hasil perhitungan daya pembeda diatas dapat disimpulkan bahwa secara keseluruhan soal PAS Pembelajaran ISMUBA Tahun ajaran 2019/2020 adalah memiliki daya pembeda yang jelek bahkan sebagaian soal bernilai tidak baik dan harus dibuang karena tidak layak.

Perhitungan efektivitas pengecoh dilakukan dengan menggunakan program Anates versi 4. Hasil perhitungan pada setiap pengecoh diinterpretasikan ke dalam lima kriteria yaitu :

IP = 76% - 125% berarti sangat baik,

IP = 51% - 75% atau 126% -150% berarti baik,

IP = 26% - 50% atau 151% - 175% berarti kurang baik,

IP = 0% - 25% atau 176% - 200% berarti jelek, dan IP = lebih dari 200% berarti sangat jelek.

Interpretasi terhadap kualitas butir soal diadaptasi dari skala *Likert* sebagai berikut (Pranatawijaya et al., 2019):

a. Apabila butir soal memenuhi empat kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut dapat dikatakan soal yang sangat baik dan dapat disimpan pada bank soal.
b. Apabila butir soal memenuhi tiga dari empat kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut dikatakan soal yang baik dan belum bisa disimpan dalam bank soal. Soal tersebut perlu direvisi sampai memenuhi empat kriteria.
c. Apabila butir soal memenuhi dua dari empat kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut dikatakan soal yang sedang dan belum bisa disimpan dalam bank soal. Soal tersebut perlu direvisi sampai memenuhi empat kriteria.
d. Apabila butir soal memenuhi satu dari empat kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut dikatakan soal yang tidak baik dan belum bisa disimpan dalam bank soal. Soal tersebut perlu direvisi secara signifikan sehingga lebih baik soal tersebut dibuang atau tidak disimpan dalam bank soal.
e. Apabila butir soal tidak memenuhi dari keempat kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut dikatakan soal yang sangat tidak baik dan

tidak bisa disimpan dalam bank soal. Soal tersebut memerlukan revisi yang signifikan sehingga lebih baik dibuang.

f. Selain syarat berbasis butir, maka tes secara keseluruhan harus reliabel dengan ketentuan yang sudah dijelaskan sebelumnya.

Berdasarkan hasil analisis, tidak terdapat butir soal yang berkualitas sangat baik, soal yang berkualitas baik berjumlah 2 butir soal, 16 butir soal termasuk dalam soal yang memiliki kualitas sedang, soal yang berkualitas tidak baik berjumlah 16 butir soal, dan terdapat 6 butir soal yang memiliki kualitas sangat tidak baik (Amalia & Widayati, 2012).

Berdasarkan hasil pengolahan data dan analisa terhadap butir soal pada Pembelajaran ISMUBA Tahun ajaran 2019/2020 (Abrori et al., 2021), didapatkan bahwa secara keseluruhan soal PAS semester ganjil tidak Reliabel dimana tidak memenuhi kriteria point yang berbunyi "*Selain syarat berbasis butir, maka tes secara keseluruhan harus reliabel dengan ketentuan yang sudah dijelaskan sebelumnya pada skala likert."* Berdasarkan distribusi perolehan kualitas soal, menjelaskan bahwa presentasi secara keseluruan atau secara umum dinyatakan tidak baik dan harus harus direvisi secara keseluruhan.

Mengacu pada kriteria yang digunakan Anderson mulai dari C1-C6 yang akan dianalisa sebagai berikut (Tondowala, 2012): Soal no. 1, 2, 4, 11, 13, Menggunakan C1 dan soalnya menunjukkan kategori yang sangat tidak baik, maka soal tersebut dibuang. Padahal pada tingkatan SMA C1 tidak diperbolehkan, karena tingkat atau level SMA sudah mencapai minimal C4. Inilah yang membuat peserta didik di Indonesia kurang kreatif dan inofatif.

Soal no. 3, 5, 6, 7, 17, 20, 21, 25, 26, 27, 28, 30, 32, 34, 35, 36, 40, menunjukkan soal yang berkategori tidak baik maka soal tersebut dibuang. Ketentuan tersebut mengharuskan bahwa soal yang digunakan memang harus kompeten dan memenuhi syarat. Jika dilihat dari kategori yang sudah dianalisa butir soalnya, terdapat 42,5% yang tidak baik. Menunjukkan kualitas guru dalam menyusun instrumen penilaian sangat rendah (Taufiqurrahman et al., 2018).

Soal no. 8, 9, 10, 14, 15, 16, 18, 19, 22, 23, 24, 31, 33, 37, 38, 39 menunjukkan soal yang berkategori sedang, maka soal tersebut diperbaharui. Ketentuan tersebut mengharuskan bahwa soal yang digunakan memang harus kompeten dan memenuhi syarat. Jika dilihat dari kategori yang sudah dianalisa butir soalnya, terdapat 40% yang sedang. Menunjukkan kualitas guru dalam menyusun instrumen penilaian belum dapat merangsang peserta didik untuk memiliki daya kritis dan kreatif (Hanifah, 2019).

Soal no. 12, dan 29 Menunjukkan soal yang berkategori baik, maka soal tersebut diterima. Ketentuan tersebut mengharuskan bahwa soal yang digunakan memang harus kompeten dan memenuhi syarat. Jika dilihat dari kategori yang sudah dianalisa butir soalnya, terdapat 5% baik. Menunjukkan kualitas guru dalam menyusun instrumen penilaian masih minim dalam menyusun instrumen penilaian. Hasil observasi yang dilakukan terhadap peserta didik menunjukkan bahwa peserta didik dalam kegiatan proses pembelajaran tidak bersikap kritis dan inovatif berdasarkan, bahwa peserta didik hanya menerima materi saja tanpa adanya keingintahuan yang mendalam. Disamping itu, peserta didik juga belum mampu membaca masalah realita seperti masih banyaknya ketidak sadaran peserta didik dalam membuang sampah, bijak memakai air serta ketidak mampuan berinofasi dalam pembelajaran di kelas. Data ini diperkuat wawancara dengan salah satu guru wali kelas SMK Muhammdiyah Mlati Ibu Risky Ulung Marsella.

"Anak-anak jika pembelajaran masih belum begitu banyak yang bertanya menganai materi yang disampaikan, terlebih lagi berpikir kritis. Tidak adanya keingintahuan peserta didik menyebabkan tumpulnya inovasi yang muncul dari akal pikran mereka. Serta ketika praktek, mereka masih tidak bisa menjawab masalah-masalah yang muncul. Seperti, ketika imam batal shalatnya, mereka tidak tahu harus melakukan apa serta mereka jarang menanyakan realita di lapangan yang terjadi" (Marsella, 2022).

Dengan demikian menunjukkan instrumen penilaian dalam mengembangkan daya kritis dan kreatif peserta didik masih tergolong sangat rendah, terbukti hanya 5% yang diterima, sedangkan 40% diperbaiki dan 55% ditolak, dibuang dan tidak diterima. Secara keseluruhan instrumen penilaian belum efektif. Maka perlu adanya evaluasi dalam pembuatan butir soal.

KESIMPULAN

Dengan demikian bisa dapat disimpulkan berdasarkan data dan hasil analisis data dari wawancara, dokumentasi, dan olah data yang mendukung penelitian serta yang berkaitan dengan pembahasan dan pengkajian mengenai kemampuan guru dalam menyusun instrumen penilaian pembelajaran ismuba meningkatkan kemampuan berfikir kritis dan kreatif peserta didik di SMK Muhammadiyah Mlati, dapat ditarik sebuah kesimpulan sesuai dengan hasil dan tujuan penilitian sebagai berikut :

1. Jika dilihat dari hasil yang didapat, maka disimpulkan bahwa kemampuan guru dalam menyusun instrumen penilaian belum sesuai dengan kompetensi jenjang SMA/MA yakni KI, KD dan Indikator. Kemampuan guru dalam menguasai materi belum sepenuhnya terlihat dengan cara penyampaian yang masih tergolong tradisionalis dan mempengaruhi daya kritis dan kreatif peserta didik.
2. Berdasarkan hasil analisis perhitungan data, tidak terdapat butir soal yang berkualitas sangat baik, soal yang berkualitas baik berjumlah 2 butir soal, 16 butir soal termasuk dalam soal yang memiliki kualitas sedang, soal yang berkualitas tidak baik berjumlah 16 butir soal, dan terdapat 6 butir soal yang memiliki kualitas sangat tidak baik.
3. Kurang efektif dengan melihat bahwa banyaknya instrumen butir soal tidak baik dan mesti dibuang serta melihat kata kata operasional yang digunakan dalam penyusunan instrument penilaian masih mengggunakan C1-C3. Maka soal atau instrumen yang digunakan belum efektif dalam meningkatkan daya kritis dan kreatif peserta didik. Selain itu, hasil observasi menunjukkan bahwa peserta didik terlihat masih rendahnya kekritisan dan kreatifan peserta didik dalam kehidupan sehari-hari.

DAFTAR PUSTAKA

Abrori, M. S., Wicaksono, Y., & Tripitasari, D. (2021). System Approach and Design Models of PAI Learning System Approach and Design Models of PAI Learning A . Pendahuluan Pendekatan sistem merupakan alat pembantu guru untuk mengambil keputusan dengan pertimbangan semua aspek permasalahan pembelajaran . Pendek. *JCIE: Journal of Contemporary Islamic Education*, *1*(2), 111–124. https://doi.org/10.25217/cie.v1i2.1589

Alwi, I. (2015). Kriteria empirik dalam menentukan ukuran sampel pada pengujian hipotesis statistika dan analisis butir. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, *2*(2). http://dx.doi.org/10.30998/formatif.v2i2.95

Amalia, A. N., & Widayati, A. (2012). Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di kota Yogyakarta tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, *10*(1). http://dx.doi.org/10.21831/jpai.v10i1.919

Fatimah, L. U., & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda dan fungsi distraktor. *AL-MANAR: Jurnal Komunikasi Dan Pendidikan Islam*, *8*(2), 37–64. https://doi.org/10.36668/jal.v8i2.115

Hanifah, N. (2019). Pengembangan instrumen penilaian Higher Order Thinking Skill (HOTS) di sekolah dasar. *Current Research in Education: Conference Series Journal*, *1*(1), 5. Google Scholar

Hartini, T. I., & Martin, M. (2020). Pengembangan instrumen soal HOTS (High Order Thinking Skill) pada mata kuliah fisika dasar 1. *JPF (Jurnal Pendidikan Fisika) Universitas Islam Negeri Alauddin Makassar*, *8*(1), 18–21. https://doi.org/10.24252/jpf.v8i1.11687

https://republika.co.id/.

Marniwati, M. (2019). Peningkatan Higher Order Thinking Skills (Hots) Melalui Strategi Pembelajaran Peningkatan Kemampuan Berpikir (Sppkb) Pada Pembelajaran Ekonomi. *Jurnal Pendidikan Tambusai*, *3*(3), 968–975. https://doi.org/10.31004/jptam.v3i3.311

Marsella, R. U. *Wawancara dengan Ibu Risky Ulung Marsella, pada tanggal 12 Juli 2022*.

Mispani, M., Abrori, M. S., Hasyim, U. A. A., Kushendar, K., & Muslimah, M. (2021). Higher Order Thinking Skills (HOTS) and Lower Order Thinking Skills (LOTS) on UM-PTKIN Questions on Islamic Religious Education Materials. *Jurnal Iqra': Kajian Ilmu Pendidikan*, *6*(2), 239–260. https://doi.org/10.25217/ji.v6i2.1702

Muluki, A. (2020). Analisis kualitas butir tes semester ganjil mata pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, *4*(1), 86–96. https://doi.org/10.23887/jisd.v4i1.23335

Pranatawijaya, V. H., Widiatry, W., Priskila, R., & Putra, P. B. A. A. (2019). Penerapan skala Likert dan skala dikotomi pada kuesioner online. *Jurnal Sains Dan Informatika*, *5*(2), 128–137. https://doi.org/10.34128/jsi.v5i2.185

Solichin, M. (2017). Analisis daya beda soal, taraf kesukaran, validitas butir tes, interpretasi hasil tes dan validitas ramalan dalam evaluasi pendidikan. *Dirasat: Jurnal Manajemen Dan Pendidikan Islam*, *2*(2), 192–213. Google Scholar

Sugiono. (2014). *Metode Penelitian Pendidikan Kualitatif, kuantitatif, dan R & D*. alfabeta. Google Scholar

Sukiman, S. (2012). Pengembangan sistem evaluasi. *Insan Madani*. Google Scholar

Suliyanto. (2017). Pelatihan Metode Pelatihan Kuantitatif. *Journal of Chemical Information and Modeling*, *5*(2), 223–232. Google Scholar

Taufiqurrahman, T., Heryandi, M. T., & Junaidi, J. (2018). Pengembangan Instrumen Penilaian Higher Order Thinking Skills Pada Mata Pelajaran Pendidikan Agama Islam. *Jurnal Pendidikan Islam Indonesia*, *2*(2), 199–206. https://doi.org/10.35316/jpii.v2i2.74

Tondowala, S. F. H. (2012). Pengembangan Tes Objektif Pilihan Ganda Berbasis Taksonomi Anderson dan Krathwohl untuk Kemampuan Membaca Bahasa Inggris Kelas VIII SMP di Kabupaten Poso, Propinsi Sulawesi Tengah. *Jurnal Pendidikan Dan Pembelajaran Bahasa Indonesia*, *1*(2). Google Scholar